The longitudinal effects of induction on beginning teachers' stress

Harmsen, Ruth; Lorenz, Michelle; Maulana, Ridwan; van Veen, Klaas

Published in:
British journal of educational psychology

DOI:
10.1111/bjep.12238

Publication date:
2019

British Journal of Educational Psychology (2019), 89, 259–287

The British Psychological Society

www.wileyonlinelibrary.com

The longitudinal effects of induction on beginning teachers' stress

Ruth Harmsen*, Michelle Helms-Lorenz, Ridwan Maulana and Klaas van Veen

Teacher Education, University of Groningen, The Netherlands

Background. Teaching is a stressful profession especially for beginning teachers. Induction programmes can support beginning teachers. Little is known concerning which elements of induction programmes can influence (the change in) teachers' stress over time.

Aims. This study aims to investigate the growth of stress causes and stress responses during the first 3 years of professional practice and to reveal the influence of induction arrangement elements on the initial level as well as the change in stress levels over the 2 years that followed.

Sample. Longitudinal data from a sample of 393 beginning teachers (56.5% female) were collected at three measurement occasions. All teachers were offered four different induction arrangement elements.

Method. Results of multiple group confirmatory factor analysis confirmed longitudinal measurement invariance. Multivariate latent growth curve modelling (MLGM) was conducted to examine the initial status, the subsequent linear growth, and the influence of the individual induction arrangement elements on the stress causes and stress responses.

Results. MLGM results show that perceived stress caused by high psychological task demands increases over time ($d = 0.22$), whereas perceived stress caused by negative pupil aspects decreases over time ($d = -0.52$). Further, workload reduction decreases the level of perceived high psychological task demands, negative social aspects, and all the stress responses. Perceived support for effective teaching behaviour decreases the level of perceived negative emotions and discontent. Further, school enculturation has an influence on the change in perceived discontent over time.

Conclusions. Perceived stress causes and stress responses can change over time. Specific induction arrangement elements appear to be powerful elements to reduce the level, and the change over time, of specific perceived stress causes and stress responses.

Despite decades of international research on teacher stress and stress interventions, many teachers still experience severe stress (Johnson & Birkeland, 2003; Skaalvik & Skaalvik, 2015). Particularly, beginning teachers seem to be more vulnerable to the pressures of the profession compared to experienced teachers (Gavish & Friedman, 2010; Goddard, O'Brien, & Goddard, 2006; Gold & Roth, 1993). Experiencing high levels of stress is

**Correspondence should be addressed to Ruth Harmsen, Teacher Education, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands (email: r.harmsen@rug.nl).*

DOI:10.1111/bjep.12238

Table 1. Induction arrangements for beginning teachers

Aspect	Aim	Example(s)
Workload reduction	To ease the job demands	No mentor tasks/less teaching hours
Supporting effective teacher behaviour	To improve the effective teacher behaviour	Lesson observation and feedback by a trained coach/mentor
Supporting school enculturation	To make the teacher familiar with the school culture/climate	Providing background information regarding the school culture/climate
Supporting professional development	To stimulate the teachers' professional development	Organizing meetings/courses for beginners

detrimental for teachers' well-being, their teaching behaviour quality, and retention (Harmsen, Helms-Lorenz, Maulana, & van Veen, 2018; Hanif, 2004). Furthermore, teacher stress harms students' achievement (Klusmann, Richter, & Lüdtke, 2016; Ronfeldt, Loeb, & Wyckoff, 2013).

Some beginning teachers are offered professional support during their first years of teaching, by means of induction programmes. An induction programme can be defined as a more or less formalized programme that aims to support beginning teachers in their first years of teaching after their pre-service education (Beijaard, Buitink, & Kessels, 2010). These programmes aim to create a smoother transition from teacher training to the first teaching job and usually include one or more of the following components (induction arrangements): (1) workload reduction (WR); (2) supporting effective teacher behaviour (TB) in the classroom; (3) supporting school enculturation (SE); and/or (4) supporting professional development (PD) (Helms-Lorenz, Slof, Vermue, & Canrinus, 2012, see Table 1).

Induction programmes are highly valued and well received by beginning teachers (Draper, O'Brien, & Christie, 2004; Hodkinson, 2006) and have shown to be beneficial for the improvement of beginning teachers' teaching quality and commitment, and significantly improve teachers' retention rates and student achievement (e.g., Helms-Lorenz, van de Grift, & Maulana, 2016; Hobson, Ashby, Malderez, & Tomlinson, 2009; Ingersoll & Strong, 2011; Smith & Ingersoll, 2004). Further, beginning teachers who receive an induction programme experience fewer stress causes compared to beginning teachers who do not receive an induction programme (Helms-Lorenz *et al.*, 2012). Therefore, it seems that induction programmes have great potential to influence the amount of stress experienced by beginning teachers. No research has investigated the influence of the individual elements of induction programmes on stress yet. Therefore, little is known about the impact of these different elements of the induction programmes on beginning teachers' stress (over time) and the question remains which elements should be included in an induction programme to reduce beginning teachers' stress. Research assessing the effectiveness of particular intervention strategies to help teachers and schools to reduce teacher stress is needed (Kyriacou, 2001) as this knowledge can help schools to provide (sufficient) targeted support for beginning teachers.

To provide a full picture of information regarding the changes in stress over time and the influence of induction arrangements on (changes in) stress causes and stress responses, it has been recommended that longitudinal studies should be undertaken (Brouwers & Tomic, 2000; Côté & Morgan, 2002; Glazerman *et al.*, 2010; Yamasaki, Sakai, & Uchida, 2006). Therefore, this study takes a 3-year longitudinal approach to exploring beginning secondary school teachers' stress and the influence of the individual induction arrangements on the development of stress over time.

Conceptual framework

Defining and modelling stress and the influence of induction

There seem to be two general perspectives on (teacher) stress. One perspective perceives stress responses (e.g., discontent) as a result of something outside of the individual, external factors (e.g., poor working conditions). There is an action (e.g., poor working condition) and a reaction (e.g., feeling discontent). In this line, Rudow (1999) and Kyriacou (2001) define teachers' stress as the experience of unpleasant, negative emotions, such as tension and anger, resulting from some aspects of their work. The other perspective hypothesizes that stress is internal; it is what goes on inside the individual as they interpret or react to what is going on around them (Gold & Roth, 1993). An example is the transactional model of stress (Folkman, 2013; Lazarus & Folkman, 1987). This model views work stress as a result of an interaction and appraisal process between the person and his or her environment. In this case, there is an action (e.g., certain working condition), an interpretation (e.g., this is a poor working condition), and a reaction (e.g., feeling discontent). Similarly, researchers conceptualize stress with both internal and external aspects: the degree of mismatch between the demands made upon an individual and the individual's ability to cope with those demands (Bakker & Demerouti, 2007; McCarthy, Lambert, Lineback, Fitchett, & Baddouh, 2015). Given the ongoing debate about and the different usages of the term 'teacher stress' in this study, the term is used more as a label indicating a specific field of (applied) research. In this study, teacher stress is divided into two components: (1) stress causes and (2) stress responses. Stress causes are the collection of aspects of the work content and the work situation influencing teachers at a cognitive, motivational, and emotional level. Stress responses are the results of the employees' mental interpretations when experiencing stress causes (Van Veldhoven, 1996).

Recently, the stress causes, high psychological task demands, negative social aspects, and negative pupil aspects, and the stress responses, tension, discontent, and negative emotions, were found to be the important stress factors to measure among BTs and therefore the focus in this study (Harmsen, Helms-Lorenz, Maulana, van Veen, & van Veldhoven, 2019). High psychological task demands concern the collection of psychological task demands that influences the teacher negatively at a cognitive, motivation, and emotional level. This concerns the pace and amount of work, education-specific workload (e.g., many hours of teaching), emotional workload, mental workload, and role conflict (see Appendix A). Negative social aspects are the stress causes concerning the negative social aspects of the teaching job, poor relationships with colleagues, a poor relationship with the supervisor, and a poor school climate. Negative pupil aspects are the stress causes related to negative pupil aspects, poor relationship with pupils and pupil misbehaviour. Discontent is a result of the employees' mental interpretations when experiencing the stress causes and concerns, lack of organizational commitment and intensions to quit (turnover). Tension is a stress response, which concerns the need for recovery, rumination, poor sleep quality, and tiredness during work. Finally, negative emotions is a stress response, which concerns emotional reactions during work (e.g., feeling sad) and lack of work pleasure.

The job demands–resources model (JD-R model; Bakker & Demerouti, 2007) is useful to frame and understand the interplay between stress causes, induction, and stress responses. The model illustrates the relationship between work characteristics, well-being, and organizational outcomes comprehensively. According to the JD-R model, two main psychological processes occur in concert: the health impairment process and the

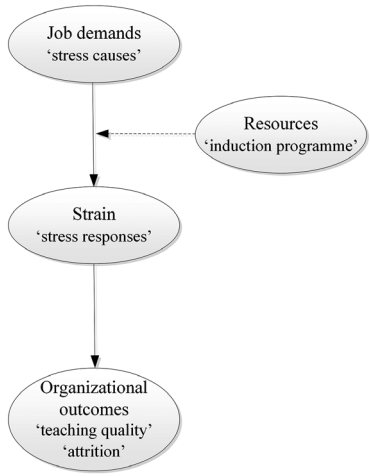


Figure 1. Proposed conceptual model of stress based on the job demands–resources model (Bakker & Demerouti, 2007).

motivational process. The first describes the relationship between job demands, job resources, strain, and organizational outcomes. The second postulates that job resources have motivational potential leading to high work engagement, low cynicism, and excellent performance. In this study, the focus is on the health impairment process (see Figure 1). Job demands are defined as 'physical, psychological, social, or organizational aspects of the job that require sustained physical and/or psychological (cognitive and emotional) effort or skills (Bakker & Demerouti, 2007)'. Translated to the teaching context, an example of this would be many hours of teaching. Job demands are not automatically stressful; however, they can turn into stress (causes) when meeting those demands requires too much effort from which the person has not adequately recovered. Hence, job demands of the JD-R model are related to stress causes as defined by Van Veldhoven (1996).

Job resources refer to physical, psychological, social, or organizational aspects of the job that are either functional in achieving work goals, reducing job demands and the associated physiological and psychological costs, or stimulate personal growth, learning, and development (Bakker & Demerouti, 2007). As can be seen in Table 1, induction programmes (see Helms-Lorenz *et al.*, 2016) provide many kinds of support for BTs and can therefore be seen as job resources.

According to the JD-R model, high job demands can exhaust employees' job resources leading to psychological strain (e.g., dissatisfaction and exhaustion; Bakker & Demerouti, 2007). Therefore, psychological strain of the JD-R model is conceptually related to stress responses as defined by Van Veldhoven (1996; negative emotions, discontent, and tension). Finally, psychological strain can lead to negative organizational outcomes such as absenteeism, poor performance, and health-related problems (Bakker, Demerouti, & Sanz-Vergel, 2014).

Several studies concerning the teaching context have found evidence for the health impairment process. Fives, Hamman, and Olivarez (2007) for example found that student–teachers who experienced high guidance during their practicum demonstrated lower levels of emotional exhaustion and depersonalization at the end of their practicum. This study shows that cooperating teachers and university supervisors can play a role in helping to reduce feelings of emotional exhaustion and depersonalization in student–

teachers. They referred to emotional exhaustion in the teaching context as the feeling of haven given all that one can, teacher has put all of his or her energy and focus into the task of teaching but finally ran out of resources. Further, depersonalization was defined as occurring when the teacher develops negative feelings and cynicism towards his or her students and perhaps even the school community. Both concepts are one of the three subconcepts of burnout, a syndrome of emotional exhaustion and cynicism that occurs frequently among individuals who do 'people-work' of some kind. The third subconcept is reduced personal accomplishment, which is the feeling of employees that they are not effective in their job (Maslach & Jackson, 1981). Gavish and Friedman (2010) found that among novice teachers, perceiving a lack of collaborative and supportive ambience contributes to predicting burnout. According to them, a collaborative supportive ambience would be a professional environment without inter- or intrarole conflict, with consistent, clear rules, and a positive and warm social climate, which is supportive of PD and which nurtures teacher competencies as an 'organizational person'.

The development of stress causes and stress responses (and the influence of induction)

A longitudinal study (four measurements during 2 years) among 79 Australian beginning teachers showed that beginning teachers experience significantly high work pressures throughout the whole measurement period and that the work pressure felt after 15 and 21 months is higher than the pressure felt after 6 weeks of teaching. Further, the study showed significant declines in the perception of job commitment, job autonomy, role clarity, co-worker cohesion, and supervisor support. They also reported a declining view of several key aspects of their work environment. In line with this finding, respondents were asked to rate their work input with respect to the level of satisfaction gained from being a teacher. At the first measurement moment, 35% of the teachers were experiencing an imbalance between the effort they were putting into their work as teachers and the rewards they received from undertaking this work, they felt that they were putting more effort into the job than they got out of the job in rewards. During the time of the study, the percentage of teachers experiencing this imbalance increased, and after 2 years, 49% of the teachers reported this imbalance. Important to rapport is that those teachers who perceived equity in the effort–reward balance reported significantly higher involvement in their work, greater support from their supervisor, and lower work pressures compared to teacher who perceived inequity (Goddard *et al.*, 2006).

Further, related to the development of stress responses over time, some studies show that beginning teachers' emotional exhaustion, depersonalization, and personal accomplishment all significantly increase during the first 2 years of teaching (Dicke *et al.*, 2015; Goddard *et al.*, 2006). However, it also has been found that emotional exhaustion decreases during the first year of teaching (Gavish & Friedman, 2010). Important to note is that teachers who perceived equity in the effort–reward transaction reported significantly lower mean scores on these variables compared to teachers who experience inequity (Goddard *et al.*, 2006). Also, professional knowledge was shown to buffer the increase in emotional exhaustion (Dicke *et al.*, 2015).

Helms-Lorenz and Maulana (2016) conducted an experimental study to investigate the influence of induction (all elements in concert) on the longitudinal relationships between self-efficacy and stress causes and stress responses. Both stress aspects were studied separately because the induction intervention aimed to reduce stress causes as well as outcomes. WR and enculturation were aimed to reduce stress causes, and support in the classroom and PD support were aimed to increase self-efficacy, which mitigates stress

responses. They revealed that the link between self-efficacy and stress responses is much stronger in the induction condition. This finding suggests that the induction arrangement (all elements in concert) is a powerful means in sustaining self-efficacy in the class that mitigates the level of job tension. This study sets out to unravel the influence of individual induction elements on the development of both stress causes and stress responses over a period of 2 years. The unanswered question is whether induction arrangement elements succeed to reduce stress causes sufficiently as the intervention might also add to high psychological task demands (by increasing accountability and performance stress) even if WR is applied.

The present study

The present study aims to answer the following research questions: (1) How do beginning teachers' perceived stress causes change across 3 years of professional practice (three measurement occasions)? Do teachers differ in their growth trajectory of perceived stress causes over time? (2) How do beginning teachers' perceived stress responses change across 3 years of professional practice? Do teachers differ in their growth trajectory of perceived stress responses over time? (3) Do induction arrangements elements provided during the first year of professional practice predict beginning teachers' experienced stress causes and stress responses at the end of the first year and over time?

Based on the existing research described above, we expect that induction arrangements influence (the change in) beginning teachers' stress causes and stress responses. We expect that beginning teachers who experience much support via induction arrangements experience in general less stress causes and stress responses (over time). As no research has investigated the influence of the individual elements of induction programmes on the individual stress causes and stress responses yet, no specific hypotheses regarding these relationships were formulated.

Method

Procedure and participants

The sample consisted of 393 (56.5% female[1]) beginning teachers from 70 secondary schools in the Netherlands. Beginning teachers in this study were defined as teachers who recently obtained their teaching qualification and who had less than 3 years' experience in the teaching profession (Kyndt, Gijbels, Grosemans, & Donche, 2016). Teachers participated in this study on a voluntary basis and received a small incentive for their participation. Informed consent was collected from all the participants.

Data for this study were collected in successive years on three measurement occasions, Time 1 (T1), Time 2 (T2), and Time 3 (T3), divided into two cohorts (see Table 2). Time 1

Table 2. Data collection timeframe

	T1	T2	T3
Cohort 1	May–June 2011	May–June 2012	March–May 2013
Cohort 2	May–June 2012	May–June 2013	March–May 2014

[1] *The percentage of female teachers at secondary schools in the Netherlands in the population is 46.7% (Ministerie van Onderwijs, Cultuur en Wetenschap, 2014).*

Table 3. Number of participants at different measurement occasions

	Frequency	Percentage female
T1	355	58.59
T2	265	56.98
T3	245	58.37

being the end of the first school year after 1 year of induction, Time 2 represents the end of the second school year after 2 years of induction, and Time 3 equals to the end of the third school year after 3 years of induction. Typically in longitudinal studies, the number of teachers per measurement occasion varies (see Table 3). Stress data were collected on all three measurement occasions (T1–T3), all three measurements were used in this study. Data regarding the perceived induction offered to the beginning teacher collected on T1 were used in this study.

Measures

Stress causes and stress responses

The stress causes and stress responses were measured at the same time with the Questionnaire on the Experience and Evaluation of Work (QEEW) (In Dutch: VBBA: Vragenlijst Beleving en Beoordeling van de Arbeid; Van Veldhoven & Meijman, 1994). This is a reliable and valid questionnaire (Evers, Van Vliet-Mulder, & Groot, 2000) developed by Van Veldhoven and Meijman (1994) who aimed to design a questionnaire that was congruent with contemporary theories of and approaches to psychosocial workload and job stress. They performed an extensive facet analysis of fifty existing instruments for measuring these issues. The models used see job stress as the result of a discrepancy between the demands of the job and the ability of the employee to meet these demands. When the job demands exceed the employees' ability to meet these demands over a longer period, health problems will arise (Van Veldhoven, Prins, van der Laken, & Dijkstra, 2015). In a recent study, the QEEW was adjusted to measure stress causes and stress responses among beginning teachers. Confirmatory factor analyses and Mokken scaling item reduction (double monotonicity Mokken model; Mokken, 1971) was applied to create high concise and precise scales. In this study, only the items and scales that proved to be of high importance to measuring stress in beginning teachers were used. There were 124 questions belonging to a factor structure with 19 first-order scales and six-second-order scales (see Appendix A; Harmsen, Helms-Lorenz, Maulana, van Veen, & van Veldhoven, 2019). It has been found that the second-order model provides a significantly better presentation of the data than the first-order model competing models (Van Veldhoven, 1996). In this study, the factor structure was retested and confirmed (see Preliminary analyses section). The second-order scales, high psychological task demands (PSY), negative social aspects (SOC), negative pupil aspects (PUPIL), tension (TEN), discontent (DIS), and negative emotions (EM), were used in further analyses in this study. Of these second-order scales, PSY, SOC, and PUPIL are stress causes, whereas TEN, DIS, and EM are stress responses.

Induction arrangements

The induction arrangements were measured using the school induction arrangement inventory (Helms-Lorenz, Slof, & van de Grift, 2013) at the end of the first induction school

year. This questionnaire measures beginning teachers' perceptions of extent to which the school undertakes supportive activities in terms of: (1) WR; (2) SE; (3) effective teaching behaviour (TB); and (4) PD. All the items were rated on a 4-point Likert scale ranging from 1 (always) to 4 (never). The item scores were recoded so that the highest scores represent much effort to support beginning teachers during induction and the lowest scores represent that schools do not provide the support to a high extent. An inventory representing each of the induction elements mentioned above was computed for use in further analyses. This inventory represents an inventory regarding the beginning teacher's perception of offered support.

Data analysis approach

Preliminary descriptive analyses, reliability analyses (using SPSS23, IBM, Armonk, NY, USA), confirmatory factor analyses (CFAs), and longitudinal invariance tests (using M*plus*, Muthén & Muthén, Los Angeles, CA, USA) were conducted. Descriptive analyses were conducted to calculate the means of the second-order stress cause and stress response scales and the means of the induction arrangement inventory. Reliability analyses were conducted to investigate the internal consistency of these scales. CFAs were conducted to examine whether the second-order factor structures, for the stress causes and stress responses, which were found in earlier research (Harmsen, Helms-Lorenz, Maulana, van Veen, & van Veldhoven, 2019), could be confirmed in this study. Longitudinal invariance tests using the multiple group confirmatory factor analysis (MGCFA) framework were conducted to establish the longitudinal invariance of the QEEW. First, univariate latent growth curve models (ULGMs) were conducted for each stress cause and stress response. The results of these ULGMs were used to create one multivariate latent growth model (MLGM) depicting the growth of stress causes and one MLGM depicting the growth of stress responses. MLGMs were conducted to examine the interindividual and intra-individual differences in stress developmental trajectories across time (using Mplus version 7.11, Muthén & Muthén, 2013). The minimum requirements of 200 participants per measurement occasion to conduct these analyses (Boomsma & Hoogland, 2001) were met (see Table 3). To evaluate the goodness of fit, commonly used fit statistics and fit indices, including the chi-square statistic, the comparative fit index (CFI), the root-mean-square error of approximation (RMSEA), and the standardized root-mean-square residual (SRMR), were used. Following the general guidelines (Byrne, 2013), the following threshold values were used to determine good and appropriate fit. Models with non-significant chi-square values, CFI and TLI values above the cut-off value of 0.95, a RMSEA value less than the cut-off value of 0.08, and SRMR less than the cut-off value of 0.08 were considered as good fit. It was noted that the chi-square statistic can easily be inflated by large sample size (Hu & Bentler, 1999).

Missing data

As is typical in longitudinal studies, and also in this study, attrition and non-response occurred over the 3-year period. 'Missingness, based on continuous data, is easily addressed in M*plus* through use of the robust maximum likelihood (MLR) estimator' (Byrne, 2013, p. 314). Therefore, the missing data in this study were handled using the MLR estimator. Also, comparative analyses with data sets including missing data ($N = 229$) and data sets without missing data (complete data; $N = 165$) were conducted.

Table 4. Comparisons among participants with missing data and complete data

	T1			T2			T3		
	N	*t*	*p*	*N*	*t*	*p*	*N*	*t*	*p*
PSY	353	−.25	.81	262	.59	.56	243	.70	.49
SOC	343	−1.63	.10	259	−.48	.65	238	.24	.81
PUPIL	353	−1.02	.31	263	−1.34	.18	243	.63	.53
DIS	**337**	**−2.86**	**.01**	253	−1.15	.25	233	.92	.36
TEN	337	−1.39	.16	252	−.10	.92	233	−.01	.99
EM	**337**	**−3.42**	**.01**	253	−.48	.63	233	.60	.48
WR	305	.23	.82				224	.34	.74
SE	304	.35	.73				224	1.00	.32
TB	301	−.96	.34				224	.56	.59
PD	301	−.36	.72				224	−1.28	.20

Note. PSY stands for high psychological task demands, SOC for negative social aspects, PUPIL for negative pupil aspects, DIS for discontent, TEN for tension, EM for negative emotions, WR for workload reduction, SE for supporting school enculturation, TB for support for effective teaching behaviour, and PD for supporting professional development. Significant results are displayed in bold.

Results showed that these data sets did not differ in terms of gender, $\chi^2_{(1,\, N\, =\, 394)} = 1.73$, $p = .19$.

For most of the scales, there were no significant differences between beginning teachers with complete data and incomplete data (see Table 4). Nevertheless, significant differences were found regarding discontent and negative emotions at T1 showing that the data set including missing data showed significantly higher scores on discontent and negative emotions at T1 compared to the data set without missing data, that is complete data. Hence, caution should be taken when interpreting the results as it seems that the more unsatisfied teachers did not respond in subsequent measures. Also, additional analyses were performed comparing stress scores of beginning teachers who left their first school with stress scores of beginning teachers who stayed at their first school.

Results

Preliminary analyses

Table 5 presents the results of the reliability and descriptive statistics analyses. These results show that the stress scales revealed a good reliability (Cronbach's alpha = .78–.94). The induction arrangement inventories had acceptable reliabilities (Cronbach's alpha = .57–.75). The stress cause, high psychological task demands, increased significantly over time ($d = 0.22$). The stress cause, negative social aspects, remained relatively stable over time. The stress cause, negative pupil aspects, decreased significantly over time ($d = -0.52$). Results also show that the stress responses, discontent, tension, and negative emotions, remained relatively stable over time. Lastly, preliminary analyses supported the adequacy to conduct CFAs. Table 6 shows the correlations among the variables considered in this study. As can be seen from this table, there are no over high dependencies among the factors. Results of the CFAs showed that the second-order factor structure, for stress causes (CFI = 0.92, TLI = 0.88, RMSEA = 0.09, SRMR = 0.06) and stress responses (CFI = 0.99, TLI = 0.98, RMSEA = 0.03, SRMR = 0.03), found in this study was acceptable and is in line with

Table 5. Descriptive statistics and reliability of the scales ($n = 393$)

		T1					T2					T3						
Stress scale	Items	Mean	Std. deviation	Minimum	Maximum	Cronbach's α	Mean	Std. deviation	Minimum	Maximum	Cronbach's α	Mean	Std. deviation	Minimum	Maximum	Cronbach's α	Effect T2 − T1	Effect T3 − T1
PSY	30	5.97	1.82	0.00	13.00	.88	6.10	1.73	2.20	12.00	.88	6.35	1.65	2.60	13.20	.87	0.07	0.22
SOC	21	3.93	1.78	0.25	9.50	.88	4.01	1.92	0.25	9.50	.90	3.93	1.82	0.00	10.00	.89	0.04	0.00
PUPIL	7	6.42	4.30	0.00	20.00	.89	4.89	3.71	0.00	19.00	.87	4.39	3.44	0.00	19.00	.85	−0.38	−0.52
DIS	12	1.72	1.36	0.00	6.00	.80	1.62	1.30	0.00	6.00	.79	1.70	1.42	0.00	6.00	.82	−0.08	−0.01
TEN	33	3.85	2.62	0.00	14.00	.94	3.85	2.27	0.00	15.25	.78	3.83	2.32	0.00	13.75	.81	0.00	−0.01
EM	17	1.07	1.38	0.00	8.50	.82	0.99	1.28	0.00	7.00	.80	1.06	1.38	0.00	8.00	.83	−0.06	−0.01
WR	17	0.60	0.19	0.00	1.00	.74												
SE	6	0.66	0.23	0.00	1.00	.57												
TB	9	0.44	0.27	0.00	1.00	.75												
PD	10	0.50	0.26	0.00	1.00	.73												

Note. PSY stands for high psychological task demands, SOC for negative social aspects, PUPIL for negative pupil aspects, DIS for discontent, TEN for tension, EM for negative emotions, WR for workload reduction, SE for supporting school enculturation, TB for support for effective teaching behaviour, and PD for supporting professional development.

Table 6. Pearson's bivariate correlations among variables considered in this study (second-order scales) (n = 393)

Scale	PSY	SOC	PUPIL	DIS	TEN	EM	WR	SE	TB	PD
PSY	1									
SOC	.55**	1								
PUPIL	.38**	.21**	1							
DIS	.33**	.46**	.15**	1						
TEN	.62**	.44**	.29**	.29**	1					
EM	.51**	.44**	.30**	.50**	.62**	1				
WR	−.25**	−.31**	.02	−.27**	−.25**	−.23**	1			
SE	−.09	−.18**	−.08	−.21**	−.06	−.16**	.40**	1		
TB	−.14*	−.18**	.05	−.21**	−.14*	−.20**	.28**	.28**	1	
PD	−.12*	−.15*	−.06	−.14**	−.09**	−.09	.33**	.45**	.44**	1

Note. PSY stands for high psychological task demands, SOC for negative social aspects, PUPIL for negative pupil aspects, DIS for discontent, TEN for tension, EM for negative emotions, WR for workload reduction, SE for supporting school enculturation, TB for support for effective teaching behaviour, and PD for supporting professional development.
*Correlation is significant at the .05 level. **Correlation is significant at the .01 level.

that found in earlier research (Appendix A; Harmsen, Helms-Lorenz, Maulana, van Veen, & van Veldhoven, 2019).

Longitudinal invariance testing

In the longitudinal term, the multiple group refers to multiple time measurement (three time points). The six-factor stress cause and stress response model was estimated simultaneously for the three time measurements. Three competing models were tested including configural, metric, and scalar invariance. The configural invariance model tests whether the six stress factors and the set of items associated with each factor have a similar factor structure over time. The metric invariance model tests whether the stress factors have the same meaning and the same measurement unit in three time measurements by assuming that factor loadings are the same across time. Achieving this level of measurement invariance (MI) allows for comparisons of latent constructs across time, but does not justify comparisons of mean scores over time. For the scalar invariance model, factor loadings (estimating the meaning of the constructs) and the intercepts are assumed to be equal over time. Reaching this level of MI allows for valid comparisons of factor scores (in our case, scale means) over time.

The common model-data goodness-of-fit indices for the categorical MGCFA model were used, including CFI, TLI, and RMSEA (Brown, 2015; Desa, 2016; Wang & Wang, 2012). Additionally, in evaluating competing models, cut-off criteria for changes in CFI (ΔCFI), TLI (ΔTLI), and RMSEA (ΔRMSEA) of <0.01 were applied (Desa, 2016; Wang & Wang, 2012).

Results of MGCFA show that the latent factor structure of stress causes and stress responses is similar over time (see Table 7, lower part), as indicated by good indices at the configural, metric, and scalar invariance (CFIs and TLIs > 0.90, RMSEAs < 0.08) as well as by the comparative assessments of relative fit indices (ΔRMSEAs, ΔCFIs, ΔTLIs < 0.01). This confirms that all indicators of stress causes and stress responses are invariant, and interpreted and responded to in a similar manner over time.

Table 7. Results of longitudinal measurement invariance testing

				RMSEA		
		CFI	TLI	Estimate	Lower bound	Upper bound
M1	Configural	0.915	0.900	0.056	0.051	0.060
M2	Metric	0.916	0.904	0.055	0.050	0.059
M3	Scalar	0.916	0.910	0.053	0.048	0.057
	Nested models comparisons	ΔCFI	ΔTLI	ΔRMSEA		
	Metric versus configural	0.001	0.004	−0.001		
	Scalar versus metric	0.000	0.006	−0.002		
	Scalar versus configural	0.001	0.010	−0.003		

Note. CFI, comparative fit index; TLI, Tucker–Lewis index; RMSEA, root-mean-square error of approximation, lower and upper for 90% confidence interval of RMSEA.
Cut-off criteria: CFI $\geq$ 0.90, TLI $\geq$ 0.90, RMSEA $<$ 0.08, ΔCFI $<$ 0.010, ΔTLI $<$ 0.010, ΔRMSEA $<$ 0.015 (Chen, 2007; Cheung & Rensvold, 2002; Desa, 2016; Wang & Wang, 2012).

Changes within and individual differences between beginning teachers' stress causes over time

To examine the within-person growth trajectory and between-person change over time concerning the beginning teachers' stress causes, ULGMs were conducted. Separate ULGMs were conducted for: high psychological task demands (PSY), negative social aspects (SOC), and negative pupil aspects (PUPIL). The trajectory of the hypothesized change is shown to be linear (no evidence of quadratic growth). Therefore, the ULGMs include two growth parameters: (1) an intercept parameter representing a beginning teacher's score on the outcome variable at Time 1, which corresponds to the end of the first school year after receiving 1 year of induction arrangements, and (2) a slope parameter representing the teachers' rate of change on the outcome variable between Time 1, Time 2, and Time 3.[2] Due to the equal interval between the measurement occasions, the time scores for the ULGMs were set as 0 for T1, 1 for T2, and 2 for T3 (see Figure 2). To take the nested nature of the data into account, the standard errors were corrected in the ULGMs using TYPE = COMPLEX.

The ULGMs depicting the growth of PSY, SOC, and PUPIL show good fit (see Appendix B). The parameter estimates from these ULGMS (see Appendix C) were used to set up the MLGMs. For the intercepts and slopes with non-significant variance, all covariances were restricted to zero in the MLGM. For the MLGM depicting the growth of stress causes, this resulted in only a covariance between the intercept and slope of SOC. Further, based on previous research (Harmsen, Helms-Lorenz, Maulana, & van Veen, 2018) and Table 6 showing correlations between PSY, SOC, and PUPIL, covariances between the intercepts of these variables were added to the MLGM. The model was run, and the structural path between the intercept and slope of SOC was not statistically significant and was therefore removed from the model to make the model more parsimonious. The resulting model, displayed in Figure 3, has a good fit; CFI = 0.907, TLI = 0.888, RMSEA = 0.078, SRMR = 0.062.

[2] *The hypothesis of linear changes was tested against the quadratic changes for all stress causes and stress responses. Results showed that no quadratic changes are evident.*

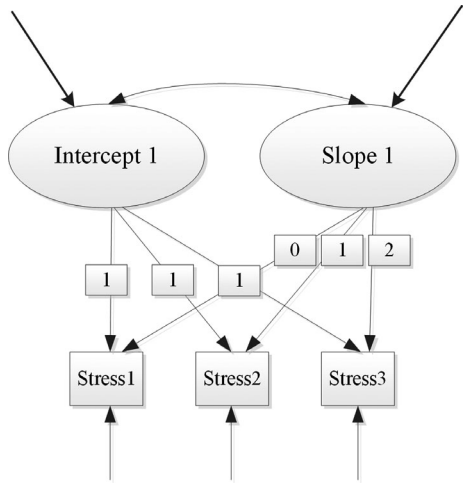


Figure 2. Univariate latent growth curve model depicting the growth of stress.

High psychological task demands

Fitting the MLGM to the teacher PSY data resulted in a mean intercept value of $M_i = 5.95$ ($p < .01$) and mean slope of $M_s = 0.25$ ($p < .01$) (see Table 8). This indicates that, in general, teachers reported a significant increase in high psychological task demands over time. The intercept variance for teacher PSY, $D_i = 2.17$ ($p < .01$), suggests that substantial variation existed between individual teachers regarding initial status of PSY. In other words, significant variation existed between individual teachers regarding PSY scores at the end of school year 1 after receiving induction arrangements for 1 year. The slope variance for PSY, $D_s = 0.12$ ($p < .05$), indicates that substantial variation existed within individual teachers regarding change in status of PSY over time. In other words, substantial variation existed within individual teachers regarding the change in PSY scores over years 1, 2, and 3.

Negative social aspects

Results show that SOC has a mean intercept value of $M_i = 3.95$ ($p < .01$) and mean slope of $M_s = 0.09$ ($p = .13$). This indicates that, in general, teachers reported no significant

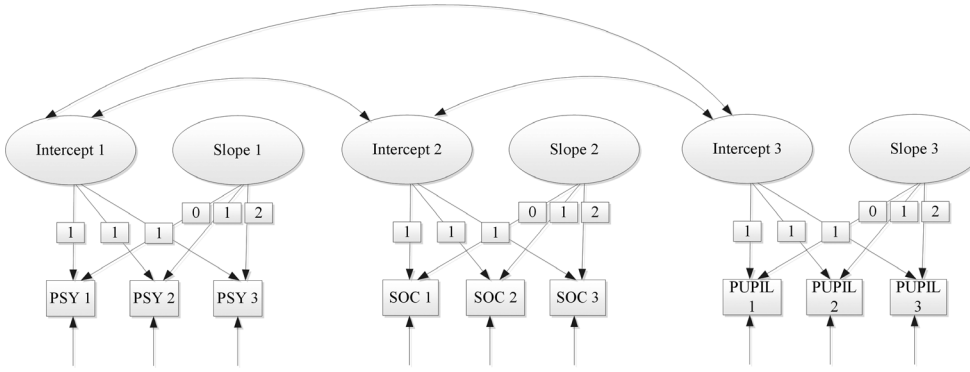


Figure 3. Multivariate latent growth curve model depicting the growth of stress causes.

Table 8. Parameter estimates for the multivariate latent growth curve model (MLGM) depicting the growth of stress causes with data set including missing data (data set 1)

Stress scale	Parameter	Estimate	*SE*	Est./*SE*	Two-tailed *p*-value
PSY	Intercept mean	**5.95**	**0.09**	**64.42**	**<.01**
	Intercept variance	**2.17**	**0.26**	**8.28**	**<.01**
	Slope mean	**0.25**	**0.04**	**6.16**	**<.01**
	Slope variance	**0.12**	**0.06**	**2.13**	**<.05**
SOC	Intercept mean	**3.95**	**0.09**	**46.53**	**<.01**
	Intercept variance	**2.20**	**0.24**	**9.28**	**<.01**
	Slope mean	0.09	0.06	1.50	n.s.
	Slope variance	**0.18**	**0.09**	**2.05**	**<.05**
PUPIL	Intercept mean	**6.21**	**0.20**	**31.16**	**<.01**
	Intercept variance	**6.71**	**1.01**	**6.68**	**<.01**
	Slope mean	**−0.90**	**0.11**	**−8.40**	**<.01**
	Slope variance	0.26	0.65	0.40	n.s.

Note. PSY stands for high psychological task demands, SOC for negative social aspects, and PUPIL for negative pupil. Significant results are displayed in bold.

change in negative social aspects over time. However, the intercept variance for teacher SOC, $D_i = 2.20$ ($p < .01$), is significant, suggesting substantial variation between individual teachers regarding SOC scores at the end of school year 1 after receiving induction arrangements for 1 year. Whereas teachers reported no significant change in negative social aspects over time, the slope variance for negative social aspects shows that there are between-teacher differences in change over time, $D_s = 0.18$ ($p < .05$).

Negative pupil aspects

Teachers reported a mean intercept value for PUPIL of $M_i = 6.21$ ($p < .01$) and a mean slope of $M_s = -0.90$ ($p < .01$). This indicates that teachers reported a significant decrease in negative pupil aspects over time. Further, results show that substantial variation existed between individual teachers regarding PUPIL scores at the end of school year 1 after receiving induction arrangements for 1 year, $D_i = 6.71$ ($p < .01$). No substantial variation existed between individual teachers regarding change in PUPIL scores over time, $D_s = 0.26$ ($p = .69$).

Changes within and individual differences between beginning teachers' stress responses over time

Separate ULGMs were also conducted to examine the growth of the stress responses: discontent (DIS), tension (TEN), and negative emotions (EM). The ULGMs depicting the growth of the stress responses show good fit (see Appendix B). The parameter estimates (see Appendix C) were used to set up the MLGMs. For the intercepts and slopes, which had non-significant variance, all covariances were restricted to zero in the MLGM, resulting in only a covariance between the intercept and slope of DIS. Further, based on previous research (Harmsen, Helms-Lorenz, Maulana, & van Veen, 2018) and Table 6 showing correlations between DIS, TEN, and EM, covariances between the intercepts of these variables were added to the MLGM. Fit statistics show that the model, displayed in

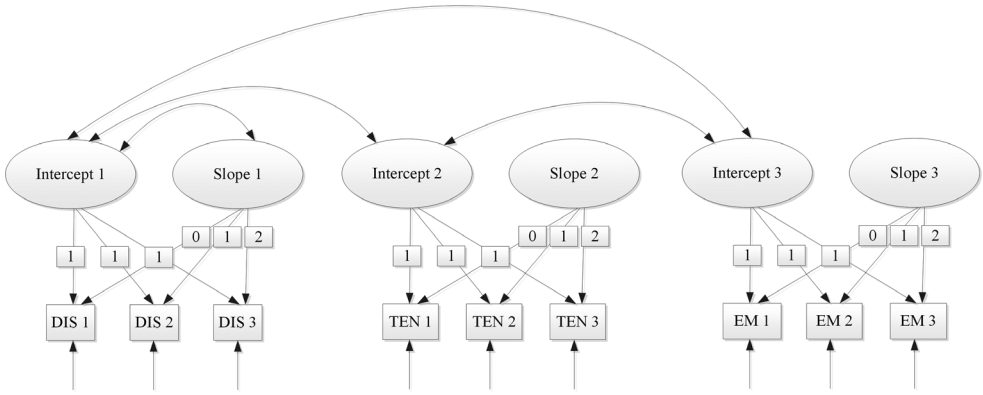


Figure 4. Multivariate latent growth curve model depicting the growth of stress responses.

Figure 4, has an appropriate fit; CFI = 0.890, TLI = 0.863, RMSEA = 0.092, SRMR = 0.060.

Tension, negative emotions, and discontent

Fitting the MLGM to the teacher stress responses data resulted in a mean intercept value for DIS of $M_i = 1.70$ ($p < .01$, see Table 9) and mean slope of $M_s = 0.11$ ($p < .05$); a mean intercept value for TEN of $M_i = 3.86$ ($p < .01$) and mean slope of $M_s = 0.09$ ($p = .15$); and a mean intercept value for EM of $M_i = 1.06$ ($p < .01$) and mean slope of $M_s = 0.07$ ($p = .15$). The results indicate a significant increase in DIS over time. Results also show that substantial variation existed between individual teachers regarding DIS, TEN, and EM scores at the end of school year 1 after receiving induction arrangements for 1 year, $D_i = 1.59$ ($p < .01$), $D_i = 4.79$ ($p < .01$), $D_i = 1.07$ ($p < .01$). The slope variance for DIS shows that there are between-teacher differences in change of DIS over time, $D_s = 0.33$ ($p < .01$).

Table 9. Parameter estimates for the multivariate latent growth curve model (MLGM) depicting the growth of stress responses with data set including missing data (data set 1)

Stress scale	Parameter	Estimate	*SE*	Est./*SE*	Two-tailed *p*-value
DIS	Intercept mean	**1.70**	**0.07**	**22.97**	**<.01**
	Intercept variance	**1.59**	**0.26**	**6.07**	**<.01**
	Slope mean	**0.11**	**0.05**	**2.04**	**<.05**
	Slope variance	**0.33**	**0.10**	**3.42**	**<.01**
TEN	Intercept mean	**3.86**	**0.13**	**28.84**	**<.01**
	Intercept variance	**4.79**	**0.44**	**10.96**	**<.01**
	Slope mean	0.09	0.06	1.45	n.s.
	Slope variance	0.13	0.12	1.12	n.s.
EM	Intercept mean	**1.06**	**0.07**	**16.08**	**<.01**
	Intercept variance	**1.07**	**0.15**	**7.23**	**<.01**
	Slope mean	0.07	0.05	1.44	n.s.
	Slope variance	−0.03	0.06	−0.61	n.s.

Note. DIS for discontent, TEN for tension, and EM for negative emotions. Significant results are displayed in bold.

Complete cases versus missing data

Significant differences were found regarding DIS and EM at T1 (see Table 4) showing that the data set including missing data (data set 1) revealed significant higher scores on discontent and negative emotions compared to the data set excluding missing values. Therefore, the MLGM analysis depicting the growth of stress causes (Figure 3), and the MLGM analysis depicting the growth of stress responses (Figure 4) were repeated, this time only including participants with complete data (data set 2), to investigate whether the results which were found using data set 1 would hold (see Appendices D and E).

When comparing the MLGMs ran with data set 1 to the MLGMs ran with data set 2, it was found that the slope of EM, $M_s = 0.14$ ($p = <0.01$), increases significantly in the MLGM when the data contain complete cases only, whereas this slope was non-significant when analysing the data including participants with missing data. Additional descriptive analyses were performed to find out more regarding the group of participants with missing data.

From the 229 participants with missing data, 65 (27%) left their first school. Of these leavers, 23 (35%) left after the first year, 30 (46%) after their second year, and 12 (19%) after their third year. The reason they left were as follows: to work at a different school (66%), to quit teaching (10%), to go travelling (3%), lack of amount of teaching hours (7%), due to poor performance (2%), private reasons (2%), and unknown (10%). An additional independent sample *t*-test was performed in order to investigate whether the beginning teachers who left their first school ($N = 65$) scored different on the stress causes and stress responses compared to the beginning teachers who stayed at their first school ($N = 328$). Beginning teachers who left their first school within 3 years scored significantly higher on all the stress causes and stress responses at T1 (see Table 10). Therefore, it seems that this group inflated the values of the stress causes and stress responses at T1.

The influence of induction arrangement elements on stress causes and stress responses

There were ULGMs conducted to examine the influence of the induction arrangement elements: WR, SE, TB, and PD on stress causes and stress responses. The results from the ULGMs (see Appendix F) were used as input for the MLGMs.

Table 10. Comparisons among participants who stayed at their first school with participants who left their first school

	T1		
	N	*t*	*p* (one-tailed)
PSY	**351**	**−2.91**	**<.01**
SOC	**342**	**−2.86**	**<.01**
PUPIL	**351**	**−2.82**	**<.01**
TEN	336	−2.30	.03
DIS	**336**	**−6.48**	**<.01**
EM	**336**	**−3.63**	**<.01**

Note. PSY stands for high psychological task demands, SOC for negative social aspects, PUPIL for negative pupil aspects, DIS for discontent, TEN for tension, and EM for negative emotions. Significant results are displayed in bold.

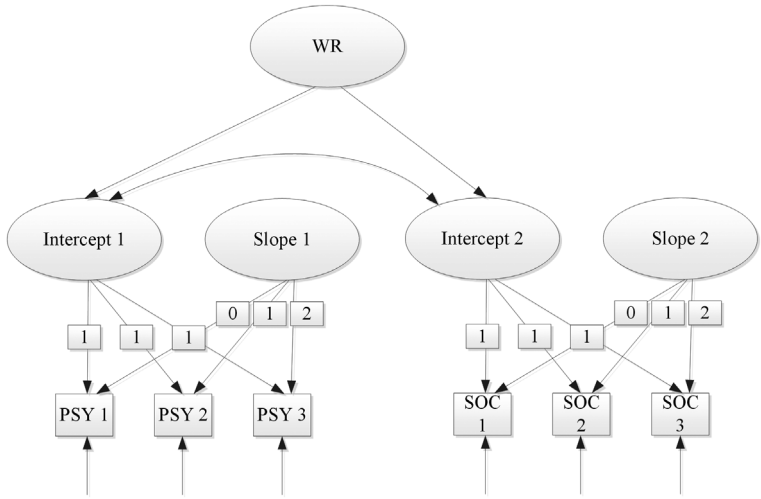


Figure 5. Multivariate latent growth curve model depicting the influence of the induction arrangement elements on stress causes. Fit statistics: CFI = 0.944, TLI = 0.927, RMSEA = 0.077, SRMR = 0.048.

For the MLGM examining the influence of induction arrangement elements on stress causes, the MLGM displayed in Figure 3 was used as the basis. Further, induction arrangements with a significant influence on the stress cause intercept or slopes (see Appendix F) were added to the MLGM. None of the induction arrangement elements significantly influence the negative pupil aspects (PUPIL). PUPIL was therefore removed from the model. By running the resulting model, it was found that the influence of TB on the intercept of SOC was no longer significant; therefore, TB was removed from the analyses. The final model (Figure 5) had a good fit; CFI = 0.944, TLI = 0.927, RMSEA = 0.077, SRMR = 0.048. The parameter estimates of the MLGM (see Table 11) show that WR has a significant negative influence on the intercept of PSY and SOC. In other words, beginning teachers who are offered WR during their first year experience less high psychological task demands and less negative social aspects at the end of that first year.

For the MLGM examining the influence of induction arrangement elements on stress responses, the MLGM displayed in Figure 4 was used as basis. Further, induction arrangements with a significant influence on the stress cause intercept or slopes (see Appendix F) were added to the MLGM. The resulting model (Figure 6) had a mediocre fit; CFI = 0.821, TLI = 0.787, RMSEA = 0.091, SRMR = 0.093. Based on the strong correlation found between WR and SE (0.40, see Table 6), a second model was run including the covariance between WR and SE. The second model had a good fit; CFI = 0.870, TLI = 0.843, RMSEA = 0.078, SRMR = 0.074.

Workload reduction

Workload reduction is a significant predictor of the initial status of PSY, SOC, TEN, EM, and DIS (see Tables 11 and 12). In other words, a high degree of perceived WR offered during the first school year corresponds to a lower degree of high psychological task demands, negative social aspects, tension, negative emotions, and discontent at the end of the first school year.

Table 11. Parameter estimates for the multivariate latent growth curve model (MLGM) depicting the influence of an induction arrangements on the intercept of stress causes

Induction arrangement	Stress scale	Parameter	Estimate	*SE*	Est./*SE*	Two-tailed *p*-value
WR	PSY	Intercept	−2.32	0.56	−4.17	<.01
WR	SOC	Intercept	−2.71	0.53	−5.08	<.01

Note. PSY stands for high psychological task demands, SOC for negative social aspects, and WR for workload reduction.

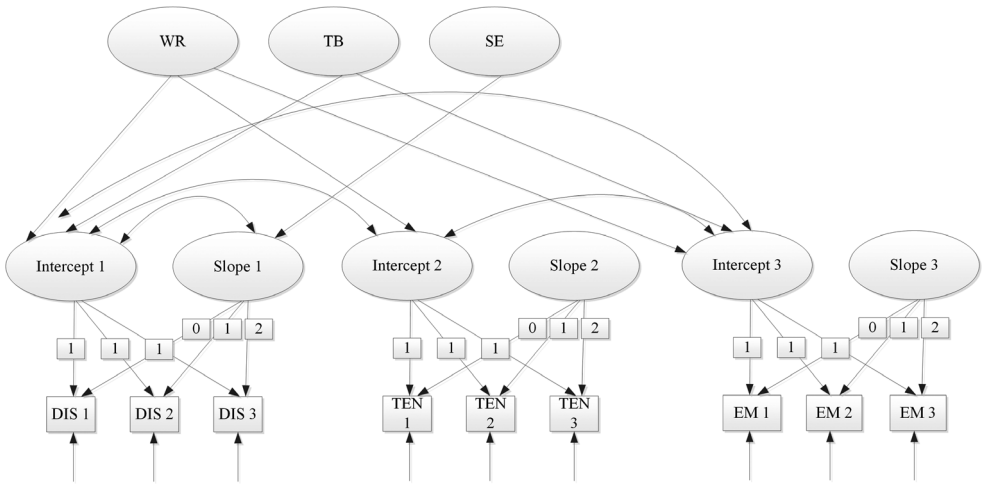


Figure 6. Multivariate latent growth curve model (MLGM) depicting the influence of the induction arrangement elements on stress responses.

Table 12. Parameter estimates for the multivariate latent growth curve model (MLGM) depicting the influence of an induction arrangements on the intercept of stress responses

Induction arrangement	Stress scale	Parameter	Estimate	*SE*	Est./*SE*	Two-tailed *p*-value
WR	DIS	Intercept	−1.39	0.34	−4.09	<.01
WR	TEN	Intercept	−2.96	0.79	−3.77	<.01
WR	EM	Intercept	−1.37	0.36	−3.87	<.01
TB	DIS	Intercept	−0.52	0.27	−1.96	0.05
TB	EM	Intercept	−0.44	0.19	−2.29	<.05
SE	DIS	Slope	−0.44	0.15	−2.94	<.01

Note. DIS stands for discontent, TEN for tension, EM for negative emotions, WR for workload reduction, TB for supporting effective teacher behaviour, and SE for stimulating school enculturation.

Stimulating SE

Stimulating SE is a significant predictor of the change in DIS. This indicates that a high degree of stimulating SE offered during the first school year corresponds with a decrease in discontent over time. However, as the reliability of SE was very low (0.57), this finding should be regarded with caution.

Supporting effective TB

Supporting effective TB is a significant predictor of the initial status of EM and DIS. This indicates that a high degree of support for effective TB offered during the first school year corresponds with a low degree of negative emotions and discontent at the end of the first school year.

Stimulating PD

Stimulating PD is not a significant predictor of the initial status nor of the change in any of the stress causes or stress responses.

Conclusions and Discussion

The present study aims to contribute to the knowledge base of teacher stress development and influences of professional support programmes by examining how beginning teachers' stress causes and stress responses develop over time and how individual induction arrangement elements provided during the first year of professional practice influence beginning teachers' experienced stress causes and responses after 1 year of teaching and over the 2 years that follow. This study shows, in line with previous research (Goddard *et al.*, 2006), that beginning teachers' perceived high psychological task demands increase significantly over time. This stress cause is perceived as most salient by the majority of beginning teachers (Borg & Riding, 1991; Boyle, Borg, Falzon, & Baglioni, 1995; Kyriacou & Sutcliffe, 1978) and has proven to be a strong predictor for feelings of tension in beginning teachers (Harmsen, Helms-Lorenz, Maulana, & van Veen, 2018). Previous research already stated that high psychological task demands and negative social aspects may be decreased by WR (Helms-Lorenz *et al.*, 2012). This study confirms that WR offered during the first year has indeed a powerful impact on decreasing the perceived high psychological demands and negative social aspects at the end of that year. That being said, none of the support offered during the first year relates to the development of this stress cause over the 2 years that followed. As this study clearly shows the benefits of WR in order to decrease the high psychological task demands, it is advised that schools keep offering this also during year 2 and year 3 to avoid work overload.

In this study, we found no significant increase concerning negative social aspects. This finding is not in line with that of Goddard *et al.* (2006) who found that co-worker cohesion and supervisor support declined over the first 21 months that beginning teachers were working. It might be that in our study, negative social aspects of the beginning teachers did not increase due to sufficient WR that was offered as we found that this influences the level of negative social aspects.

Warm teacher–student relationships are often mentioned as one of the main reasons for teachers to stay in the profession (O'Connor, 2008; Wilhelm, Dewhurst-Savellis, & Parker, 2000), whereas poor teacher–student relationships are strongly related to experiencing tension, discontent, and negative emotions (Harmsen, Helms-Lorenz, Maulana, & van Veen, 2018; Yoon, 2002). This study shows that teachers' perceived negative pupil aspects decreased significantly over time. The induction arrangements offered during the first year had no significant influence on the stress cause, perceived negative pupil aspects at the end of that year, or on the decrease in this stress aspect over time. This is an interesting finding. If the induction arrangement elements do not influence

the decrease in perceived negative pupil aspects of beginning teachers what does? In this study, the focus was only on resources, which were offered by the school, so-called contextual resources, whereas 'individual resources' such as active coping skills influence the stress process of teachers too (Beltman, Mansfield, & Price, 2011). It might be that beginning teachers who stay in the profession develop active coping skills, which helps them to cope efficiently with negative pupil aspects. Also, it might be that as teachers become more competent in teaching skills such as classroom management, pupil misbehaviour decreases. Further, building warm teacher–student relationships might be a skill which teachers develop naturally over time. Future research will benefit from investigating the relationship between the development of teacher–student relationships, teaching skills (e.g., classroom management), and (individual) resources and teacher stress over time.

We found that discontent increased significantly over time. This is in line with prior research showing that feelings of emotional exhaustion, depersonalization, and personal accomplishment all significantly increase during the first 2 years of teaching (Dicke *et al.*, 2015; Goddard *et al.*, 2006). It is certainly worth paying attention to this finding, because feeling discontent has been found to be related to beginning teachers' attrition (Harmsen, Helms-Lorenz, Maulana, & van Veen, 2018). In the light of retaining beginning teachers, it is important to highlight that we found that WR and SE can both decrease the feelings of discontent. Further, SE offered in the first year can influence the change in discontent over the 2 years that follow. Feelings of discontent decrease over time when beginning teachers experience SE. This highlights the importance of making beginning teachers familiar with the school culture and school climate. However, due to low reliability of the SE scale, this result should be regarded with caution.

The induction element that supports PD does not influence any of the stress causes or stress responses in this study. This result surprises us as previous research found that an environment that lacks support for PD contributes to predicting burnout (Gavish & Friedman, 2010). Much research is needed to explain this finding.

This study has a number of limitations. First, attrition and non-response occurred over the 2-year period influencing the results of discontent and negative emotions. Therefore, caution should be taken when interpreting the results of discontent and negative emotions in this study. Also, participants participated on a voluntary basis; therefore, it is possible that beginning teachers who agreed to participate may behave differently from those who did not. In the light of experiencing stress, it might be that severely stressed teachers did not find the time to participate. Another limitation is that the support from induction arrangements was not measured at every time point in which stress was measured. Measuring the induction arrangements at all three time points could have provided us with more insight into the more powerful influence of the arrangements per year. Finally, the variables were assessed via self-reports, which can have several limitations (Trenberth & Dewe, 2004).

Nevertheless, longitudinal studies concerning the development of stress among beginning teachers are scare; therefore, this study is an important contribution to the scientific knowledge concerning the longitudinal development of stress among beginning teachers. Also, our contribution to the field is that evidence has been offered showing that induction elements not only reduce stress causes but also have an impact on stress responses too. Beginning teachers who are supported with induction arrangements have reduced levels of discontent and reduced levels of negative emotions. Finally, as mentioned in the introduction, no research has investigated the influence of the individual induction arrangement elements yet. This study highlights which induction arrangements

elements reduce the level of beginning teachers' stress causes and stress responses (over time). This important knowledge can help schools target their interventions more precisely after diagnosing the stress causes and stress responses of their beginning teachers. Much research, particularly in-depth interviews with teachers, is needed to shed more light on explaining the complex results related to teacher stress found in this study, such as the decrease in perceived negative pupil aspects over time.

Acknowledgements

This study was part of the PhD project of the first author, while the second and the third author received a grant from the Dutch Ministry of Education under Grant (OND/ODB-13/19888). Copyrights for the use of the QEEW (translated title of the 'Vragenlijst Beleving en Beoordeling van de Arbeid' VBBA © 2015) belong to SKB, www.skb.nl. The licence to use the questionnaire was obtained. The first author would also like to thank Peter Moorer for his support during the revision of this manuscript.

References

Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of Managerial Psychology*, *22*, 309–328. https://doi.org/10.1108/01425450710826122

Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. I. (2014). Burnout and work engagement: The JD–R approach. *Annual Review of Organizational Psychology*, *1*, 389–411. https://doi.org/10.1146/annurev-orgpsych-031413-091235

Beijaard, D., Buitink, J., & Kessels, C. (2010). Teacher induction. In B. McGaw, P. L. Peterson, & E. Baker (Eds.), *International encyclopedia of education* (Vol. 7, pp. 563–568). Oxford, UK: Elsevier Scientific Publishers.

Beltman, S., Mansfield, C., & Price, A. (2011). Thriving not just surviving: A review of research on teacher resilience. *Educational Research Review*, *6*, 185–207. https://doi.org/10.1016/j.edurev.2011.09.001

Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. du Toit, & D. Sorbom (Eds.), *Structural equation models: Present and future. A Festschrift in Honor of Karl Jöreskog* (Vol. 2, pp. 139–168). Chicago, IL: Scientific Software International. Retrieved from https://core.ac.uk/download/pdf/20770852.pdf

Borg, M. G., & Riding, R. J. (1991). Towards a model for the determinants of occupational stress among schoolteachers. *European Journal of Psychology of Education*, *6*, 355–373. https://doi.org/10.1007/BF03172771

Boyle, G. J., Borg, M. G., Falzon, J. M., & Baglioni, A. J. (1995). A structural model of the dimensions of teacher stress. *British Journal of Educational Psychology*, *65*, 49–67. https://doi.org/10.1111/j.2044-8279.1995.tb01130.x

Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education*, *16*, 239–253. https://doi.org/10.1016/S0742-051X(99)00057-8

Brown, T. A. (2015). CFA with equality constraints, multiple groups, and mean structures. In T. A. Brown (Ed.), *Confirmatory factor analysis for applied research* (2nd ed., pp. 236–319). New York, NY: The Guilford Press.

Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. London, UK: Routledge.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*, 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modelling*, *9*, 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Côté, S., & Morgan, L. M. (2002). A longitudinal analysis of the association between emotion regulation, job satisfaction, and intentions to quit. *Journal of Organizational Behaviour*, *23*, 947–962. https://doi.org/10.1002/job.174

Desa, D. (2016). Understanding non-linear modeling of measurement invariance in heterogeneous populations. *Advances in Data Analysis and Classification*, *1*, 1–25. https://doi.org/10.1007/s11634-016-0240-3

Dicke, T., Parker, P. D., Holzberger, D., Kunina-Habenicht, O., Kunter, M., & Leutner, D. (2015). Beginning teachers' efficacy and emotional exhaustion: Latent changes, reciprocity, and the influence of professional knowledge. *Contemporary Educational Psychology*, *41*, 62–72. https://doi.org/10.1016/j.cedpsych.2014.11.003

Draper, J., O'Brien, J., & Christie, F. (2004). First impressions: The new teacher induction arrangements in Scotland. *Journal of In-service Education*, *30*, 201–224. https://doi.org/10.1080/13674580100200242

Evers, A., Van Vliet-Mulder, J. C., & Groot, C. J. (2000). Documentatie van tests en testresearch in Nederland [Documentation of tests and test research in the Netherlands]. In A. Evers, & C. J. Groot (Eds.), *Testbeschrijvingen* [*Test descriptions*] (pp. 649–650). Assen, The Netherlands: Van Gorcum & Comp. B.V.

Fives, H., Hamman, D., & Olivarez, A. (2007). Does burnout begin with student-teaching? Analyzing efficacy, burnout, and support during the student-teaching semester. *Teaching and Teacher Education*, *23*, 916–934. https://doi.org/10.1016/j.tate.2006.03.013

Folkman, S. (2013). Stress: Appraisal and coping. In M. D. Gellman (Ed.), *Encyclopedia of behavioral medicine* (pp. 1913–1915). New York, NY: Springer.

Gavish, B., & Friedman, I. A. (2010). Novice teachers' experience of teaching: A dynamic aspect of burnout. *Social Psychology of Education*, *13*, 141–167. https://doi.org/10.1007/s11218-009-9108-0

Glazerman, S., Isenberg, E., Dolfin, S., Bleeker, M., Johnson, A., Grider, M., & Jacobus, M. (2010). *Impacts of comprehensive teacher induction: Final results from a randomized controlled study* (No. 691d9603eb074051b57684e4affae4d4). Mathematica Policy Research. Retrieved from https://www.mathematica-mpr.com/-/media/publications/pdfs/education/teacherinduction-fnlrpt.pdf

Goddard, R., O'Brien, P., & Goddard, M. (2006). Work environment predictors of beginning teacher burnout. *British Educational Research Journal*, *32*, 857–874. https://doi.org/10.1080/01411920600989511

Gold, Y., & Roth, R. A. (1993). *Teachers managing stress and preventing burnout: The professional health solution*. London, UK: The Falmer Press.

Hanif, R. (2004). *Teacher stress, job performance and self-efficacy of women school teachers*. Doctoral thesis, National Institute of Psychology, Quaid-e-Azam University, Islamabad, Pakistan.

Harmsen, R., Helms-Lorenz, M., Maulana, R., & van Veen, K. (2018). The relationship between beginning teachers' stress causes, stress responses, teaching behaviour and attrition. *Teachers and Teaching*, *24*(6), 626–643. https://doi.org/10.1080/13540602.2018.1465404

Harmsen, R., Helms-Lorenz, M., Maulana, R., van Veen, K., & van Veldhoven, M. (2019). Measuring general and specific stress causes and stress responses among beginning secondary school teachers in the Netherlands. *International Journal of Research & Method in Education*, *42*, 91–108. https://doi.org/10.1080/1743727X.2018.1462313

Helms-Lorenz, M., & Maulana, R. (2016). Influencing the psychological well-being of beginning teachers across three years of teaching: Self-efficacy, stress causes, job tension and job discontent. *Educational Psychology*, *36*(3), 569–594. https://doi.org/10.1080/01443410.2015.1008403

Helms-Lorenz, M., Slof, B., & van de Grift, W. (2013). First year effects of induction arrangements on beginning teachers' psychological processes. *European Journal of Psychology of Education*, *28*(4), 1265–1287. https://doi.org/10.1007/s10212-012-0165-y

Helms-Lorenz, M., Slof, B., Vermue, C. E., & Canrinus, E. T. (2012). Beginning teachers' self-efficacy and stress and the supposed effects of induction arrangements. *Educational Studies*, *38*(2), 189–207. https://doi.org/10.1080/03055698.2011.598679

Helms-Lorenz, M., van de Grift, W., & Maulana, R. (2016). Longitudinal effects of induction on teaching skills and attrition rates of beginning teachers. *School Effectiveness and School Improvement*, *27*(2), 178–204. https://doi.org/10.1080/09243453.2015.1035731

Hobson, A. J., Ashby, P., Malderez, A., & Tomlinson, P. D. (2009). Mentoring beginning teachers: What we know and what we don't. *Teaching and Teacher Education*, *25*, 207–216. https://doi.org/10.1016/j.tate.2008.09.001

Hodkinson, A. (2006). Career entry development profiles and the statutory induction arrangements in England: A model of effective practice for the professional development of Newly Qualified Teachers? *Journal of In-service Education*, *32*, 287–300. https://doi.org/10.1080/13674580600841612

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55. https://doi.org/10.1080/10705519909540118

Ingersoll, R. M., & Strong, M. (2011). The impact of induction and mentoring programs for beginning teachers: A critical review of the research. *Review of Educational Research*, *81*, 201–233. https://doi.org/10.3102/0034654311403323

Johnson, S. M., & Birkeland, S. E. (2003). Pursuing a "sense of success": New teachers explain their career decisions. *American Educational Research Journal*, *40*, 581–617. https://doi.org/10.3102/00028312040003581

Klusmann, U., Richter, D., & Lüdtke, O. (2016). Teachers' emotional exhaustion is negatively related to students' achievement: Evidence from a large-scale assessment study. *Journal of Educational Psychology*, *108*, 1193. https://doi.org/10.1037/edu0000125

Kyndt, E., Gijbels, D., Grosemans, I., & Donche, V. (2016). Teachers everyday professional development: Mapping informal learning activities, antecedents, and learning outcomes. *Review of Educational Research*, *86*, 1111–1150. https://doi.org/10.3102/0034654315627864

Kyriacou, C. (2001). Teacher stress: Directions for future research. *Educational Review*, *53*, 27–35. https://doi.org/10.1080/00131910120033628

Kyriacou, C., & Sutcliffe, J. (1978). Teacher stress: Prevalence, sources, and symptoms. *British Journal of Educational Psychology*, *48*, 159–167. https://doi.org/10.1111/j.2044-8279.1978.tb02381.x

Lazarus, R. S., & Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of Personality*, *1*, 141–169. https://doi.org/10.1002/per.2410010304

Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Organizational Behavior*, *2*, 99–113. https://doi.org/3000281

McCarthy, C. J., Lambert, R. G., Lineback, S., Fitchett, P., & Baddouh, P. G. (2015). Assessing teacher appraisals and stress in the classroom: Review of the classroom appraisal of resources and demands. *Educational Psychology Review*, *28*, 577–603. https://doi.org/10.1007/s10648-015-9322-6

Ministerie van Onderwijs, Cultuur en Wetenschap [Dutch Ministry of Education] (2014). *Kerncijfers 2009-2013, Onderwijs, Cultuur en Wetenschap [Key Figures 2009-2013, Education, Culture and Science]*. Retrieved from https://www.rijksoverheid.nl/documenten/jaarverslagen/2014/05/21/ocw-kerncijfers

Mokken, R. J. (1971). *A theory and procedure of scale analysis*. The Hague, The Netherlands: Mouton.

Muthén, L. K., & Muthén, B. O. (2013). *Mplus 7.11*. Los Angeles, CA: Muthén & Muthén.

O'Connor, K. E. (2008). "You choose to care": Teachers, emotions and professional identity. *Teaching and Teacher Education*, *24*, 117–126. https://doi.org/10.1016/j.tate.2006.11.008

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, *50*, 4–36. https://doi.org/10.3102/0002831212463813

Rudow, B. (1999). *Stress and burnout in the teaching profession: European studies, issues, and research perspectives*. New York, NY: Cambridge University Press.

Skaalvik, E. M., & Skaalvik, S. (2015). Job satisfaction, stress and coping strategies in the teaching profession – What do teachers say? *International Education Studies*, *8*, 181–192. https://doi.org/10.5539/ies.v8n3p181

Smith, T. M., & Ingersoll, R. M. (2004). What are the effects of induction and mentoring on beginning teacher turnover? *American Educational Research Journal*, *41*, 681–714. https://doi.org/10.3102/00028312041003681

Trenberth, L., & Dewe, P. (2004). Work stress and coping: Drawing together research and practice. *British Journal of Guidance & Counselling*, *32*, 143–156. https://doi.org/10.1080/03069880410001692201

Van Veldhoven, M. (1996). *Psychosociale arbeidsbelasting en werkstress* [Psycho-social workpressure and workstress]. Doctoral thesis, Rijksuniversiteit Groningen, Groningen, The Netherlands.

Van Veldhoven, M., & Meijman, T. F. (1994). *Het meten van psychosociale arbeidsbelasting met een vragenlijst: de vragenlijst beleving en beoordeling van de arbeid (VBBA) [Measuring psycho-social workpressure and workstress with a questionnaire: The Questionnaire on the Experience and Evaluation of Work (QEEW)]*. Amsterdam, The Netherlands: Nederlands Instituut voor Arbeidsomstandigheden.

Van Veldhoven, M. J. P. M., Prins, J., van der Laken, P. A., & Dijkstra, L. (2015). *QEEW2. 0: 42 short scales for survey research on work, well-being and performance*. Amsterdam, Netherlands: SKB.

Wang, Z., & Wang, X. (2012). *Structural equation modelling: Applications using Mplus*. Chichester, UK: John Wiley & Sons Ltd.

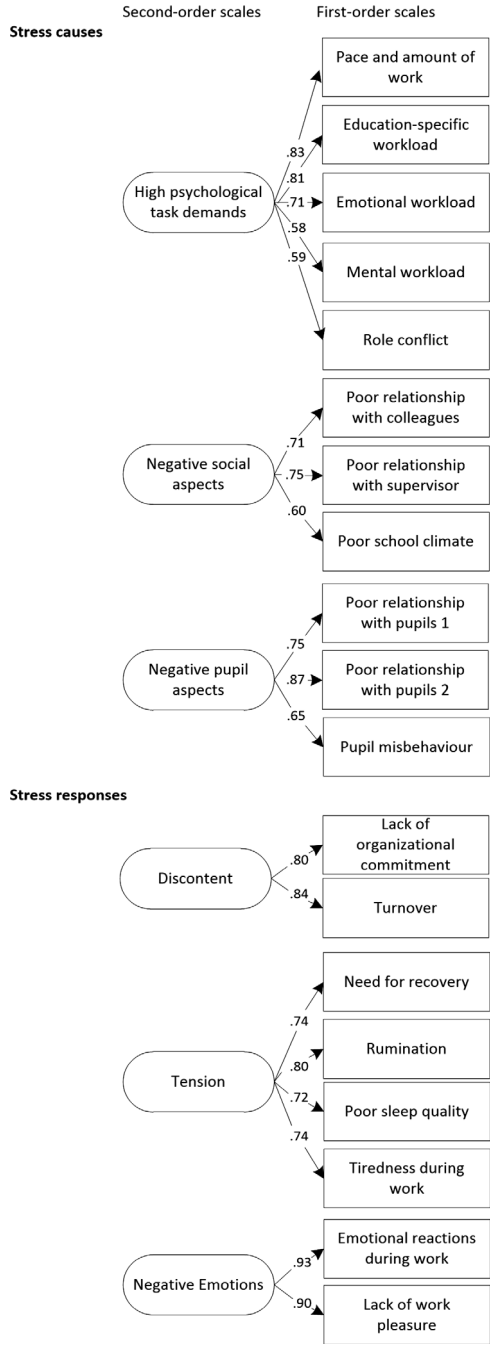
Wilhelm, K., Dewhurst-Savellis, J., & Parker, G. (2000). Teacher stress? An analysis of why teachers leave and why they stay. *Teachers and Teaching*, *6*, 291–304. https://doi.org/10.1080/713698734

Yamasaki, K., Sakai, A., & Uchida, K. (2006). A longitudinal study of the relationship between positive affect and both problem-and emotion focused coping strategies. *Social Behaviour and Personality: An International Journal*, *34*, 499–510. https://doi.org/10.2224/sbp.2006.34.5.499

Yoon, J. (2002). Teacher characteristics as predictors of teacher–student relationships: Stress, negative affect, and self-efficacy. *Social Behavior and Personality: An International Journal*, *30*, 485–493. https://doi.org/10.2224/sbp.2002.30.5.485

Received 20 July 2017; revised version received 13 June 2018

Appendix A: Factor structure stress causes and stress responses



Appendix B: Model fit statistics and fit indices of the univariate latent growth models (ULGMs) depicting the growth of stress

	Stress causes			Stress responses		
Fit statistics	PSY	SOC	PUPIL	DIS	TEN	EM
Chi-square						
Value	0.579	1.180	3.504	0.843	0.547	0.304
Degrees of freedom	1	1	1	1	1	1
p-Value	.447	.277	.061	.359	.460	.582
Scaling correction factor for MLR	0.999	0.857	1.673	1.251	1.349	1.367
CFI	1.00	1.00	0.95	1.00	1.00	1.00
TLI	1.01	1.00	0.85	1.00	1.01	1.04
RMSEA						
Estimate	0.00	0.02	0.08	0.00	0.00	.00
Confidence interval	0.000, 0.121	0.000, 0.139	0.000, 0.177	0.000, 0.132	0.000, 0.122	0.000, 0.111
SRMR	0.008	0.011	0.034	0.011	0.009	0.009

Note. PSY stands for high psychological task demands, SOC for negative social aspects, PUPIL for negative pupil aspects, DIS for discontent, TEN for tension, and EM for negative emotions.

Appendix C: Parameter estimates for the univariate latent growth curve models (ULGMs) depicting the growth of stress

Stress scale	Parameter	Estimate	*SE*	Est./*SE*	Two-tailed *p*-value
PSY	Intercept mean	**5.95**	**0.09**	**64.44**	**<.01**
	Intercept variance	**2.17**	**0.32**	**6.85**	**<.01**
	Slope mean	**0.24**	**0.04**	**5.76**	**<.01**
	Slope variance	0.08	0.15	0.54	n.s.
SOC	Intercept mean	**3.94**	**0.08**	**47.63**	**<.01**
	Intercept variance	**2.41**	**0.36**	**6.75**	**<.01**
	Slope mean	0.08	0.06	1.31	n.s.
	Slope variance	**0.43**	**0.16**	**2.59**	**<.01**
PUPIL	Intercept mean	**6.22**	**0.20**	**30.43**	**<.01**
	Intercept variance	**7.58**	**2.54**	**2.98**	**<.01**
	Slope mean	**−0.94**	**0.10**	**−9.14**	**<.01**
	Slope variance	0.47	1.39	0.34	n.s.
DIS	Intercept mean	**1.70**	**0.08**	**22.83**	**<.01**
	Intercept variance	**1.45**	**0.27**	**5.32**	**<.01**
	Slope mean	0.10	0.06	1.66	n.s.

Continued

Appendix C. (*Continued*)

Stress scale	Parameter	Estimate	*SE*	Est./*SE*	Two-tailed *p*-value
	Slope variance	**0.32**	**0.11**	**2.83**	**<.01**
TEN	Intercept mean	**3.86**	**0.13**	**28.94**	**<.01**
	Intercept variance	**4.56**	**0.70**	**6.56**	**<.01**
	Slope mean	0.04	0.07	0.63	n.s.
	Slope variance	0.26	0.40	0.65	n.s.
EM	Intercept mean	**1.06**	**0.07**	**16.18**	**<.01**
	Intercept variance	**0.89**	**0.28**	**3.17**	**<.01**
	Slope mean	0.06	0.06	0.98	n.s.
	Slope variance	−0.08	0.16	−0.50	n.s.

Note. PSY stands for high psychological task demands, SOC for negative social aspects, PUPIL for negative pupil aspects, DIS for discontent, TEN for tension, and EM for negative emotions. Significant results are displayed in bold.

Appendix D: Parameter estimates for the multivariate latent growth curve model (MLGM) depicting the growth of stress causes with complete cases only (data set 2)

Stress scale	Parameter	Estimate	*SE*	Est./*SE*	Two-tailed *p*-value
PSY	Intercept mean	**5.85**	**0.11**	**54.91**	**<.01**
	Intercept variance	**1.73**	**0.25**	**7.02**	**<.01**
	Slope mean	**0.24**	**0.04**	**5.48**	**<.01**
	Slope variance	0.18	0.07	2.62	**<.01**
SOC	Intercept mean	**3.77**	**0.11**	**34.69**	**<.01**
	Intercept variance	**1.74**	**0.24**	**7.35**	**<.01**
	Slope mean	0.09	0.07	1.35	0.18
	Slope variance	**0.21**	**0.08**	**2.45**	**<.05**
PUPIL	Intercept mean	**5.92**	**0.23**	**25.71**	**<.01**
	Intercept variance	**5.54**	**0.86**	**6.45**	**<.01**
	Slope mean	**−0.87**	**0.13**	**−6.82**	**<.01**
	Slope variance	0.38	0.69	0.55	0.59

Note. PSY stands for high psychological task demands, SOC for negative social aspects, and PUPIL for negative pupil. Significant results are displayed in bold.
Fit statistics: CFI = 0.921, TLI = 0.905, RMSEA = 0.091, SRMR = 0.053.

Appendix E: Parameter estimates for the multivariate latent growth curve model (MLGM) depicting the growth of stress responses with complete cases only (data set 2)

Stress scale	Parameter	Estimate	*SE*	Est./*SE*	Two-tailed *p*-value
DIS	Intercept mean	**1.45**	**0.08**	**17.41**	**<.01**
	Intercept variance	**1.12**	**0.25**	**4.54**	**<.01**
	Slope mean	**0.12**	**0.05**	**2.27**	**<.05**
	Slope variance	**0.29**	**0.09**	**3.33**	**<.01**
TEN	Intercept mean	**3.72**	**0.19**	**19.99**	**<.01**
	Intercept variance	**4.03**	**0.58**	**6.92**	**<.01**
	Slope mean	0.09	0.07	1.36	n.s.
	Slope variance	0.19	0.11	1.66	n.s.
EM	Intercept mean	**0.80**	**0.07**	**11.49**	**<.01**
	Intercept variance	**0.68**	**0.15**	**4.60**	**<.01**
	Slope mean	**0.14**	**0.05**	**2.88**	**<.01**
	Slope variance	−0.02	.06	−0.30	n.s.

Note. DIS for discontent, TEN for tension, and EM for negative emotions. Significant results are displayed in bold.
Fit statistics: CFI = 0.881, TLI = 0.853, RMSEA = 0.120, SRMR = 0.078.

Appendix F: Parameter estimates for the univariate latent growth curve models (ULGMs) depicting the influence of the induction arrangements on the intercept (and slope) of stress

ULGM	Induction arrangement	Parameter	Estimate	*SE*	Est./*SE*	Two-tailed *p*-value
PSY	WR	Intercept	**−2.19**	**0.61**	**−3.62**	**<.01**
	SE	Intercept	0.46	0.47	0.97	n.s.
	PD	Intercept	−0.13	0.44	−0.29	n.s.
	TB	Intercept	−0.52	0.37	−1.42	n.s.
SOC	WR	Intercept	**−1.98**	**0.50**	**−3.96**	**<.01**
	SE	Intercept	−0.25	0.48	−0.53	n.s.
	PD	Intercept	−0.11	0.51	−0.22	n.s.
	TB	Intercept	**−0.82**	**0.41**	**−1.98**	**<.05**
	WR	Slope	−0.37	0.29	−1.28	n.s.
	SE	Slope	−0.11	0.30	−0.37	n.s.
	PD	Slope	0.25	0.30	0.85	n.s.
	TB	Slope	0.24	0.27	0.89	n.s.
PUPIL	WR	Intercept	1.21	1.09	1.12	n.s.
	SE	Intercept	−1.46	0.78	−1.86	n.s.
	PD	Intercept	−0.91	0.85	−1.08	n.s.

Continued

Appendix F. (*Continued*)

ULGM	Induction arrangement	Parameter	Estimate	*SE*	Est./*SE*	Two-tailed *p*-value
	TB	Intercept	0.68	0.75	0.91	n.s.
DIS	WR	Intercept	**−1.43**	**0.44**	**−3.24**	**<.01**
	SE	Intercept	−0.02	0.47	−0.04	n.s.
	PD	Intercept	0.10	0.42	0.23	n.s.
	TB	Intercept	**−0.79**	**0.28**	**−2.82**	**<.01**
	WR	Slope	0.00	0.29	0.01	n.s.
	SE	Slope	**−0.57**	**0.26**	**−2.24**	**<.05**
	PD	Slope	0.16	0.20	0.77	n.s.
	TB	Slope	0.16	0.23	0.68	n.s.
TEN	WR	Intercept	**−3.07**	**0.81**	**−3.82**	**<.01**
	SE	Intercept	0.60	0.58	1.02	n.s.
	PD	Intercept	0.10	0.73	0.14	n.s.
	TB	Intercept	−0.78	0.46	−1.70	n.s.
EM	WR	Intercept	**−1.17**	**0.40**	**−2.97**	**<.01**
	SE	Intercept	−0.38	0.38	−1.00	n.s.
	PD	Intercept	0.39	0.36	1.08	n.s.
	TB	Intercept	**−0.71**	**0.27**	**−2.65**	**<.01**

Note. PSY stands for high psychological task demands, SOC for negative social aspects, PUPIL for negative pupil aspects, DIS for discontent, TEN for tension, EM for negative emotions, WR for workload reduction, SE for supporting school enculturation, TB for support for effective teaching behaviour, and PD for supporting professional development. Significant results are displayed in bold.